AIP | Conference Proceedings

Recent advances on the parallelization of Gauss methods

Luigi Brugnano, Felice Iavernaro, and Tiziana Susca

Citation: AIP Conf. Proc. **1479**, 238 (2012); doi: 10.1063/1.4756107

View online: http://dx.doi.org/10.1063/1.4756107

View Table of Contents: http://proceedings.aip.org/dbt/dbt.jsp?KEY=APCPCS&Volume=1479&Issue=1

Published by the American Institute of Physics.

AIPAdvances
Submit Now
Explore AIP's new open-access journal
• Article-level metrics now available
• Join the conversation! Rate & comment on articles

Recent Advances on the Parallelization of Gauss Methods

Luigi Brugnano[*], Felice Iavernaro[†] and Tiziana Susca[†]

[*]*Dipartimento di Matematica "U.Dini", University of Firenze, Italy*
[†]*Dipartimento di Matematica, University of Bari, Italy*

Abstract. We introduce a new formulation of Gauss collocation methods for the numerical solution of ordinary differential equations. These formulae may be thought of as Runge-Kutta methods with rank-deficient array and may be specified in order to allow an easy parallel implementation. We show some preliminary results on Gauss methods of order 4, 6 and 8.

Keywords: Gauss method, collocation Runge-Kutta methods, nonlinear splitting, symplectic integrators.
PACS: 02.60.-x,45.20.dh, 45.20.Jj, 02.30.Hq, 02.30.Nw.
AMS MSC: 65P10, 65L05, 65L06.

INTRODUCTION

Gauss collocation methods play a prominent role in the numerical treatment of Initial Value Problems (IVPs) and of more general differential equations. Its classical convergence and stability properties (see [7]) are very well known to any researcher in this field. However, due to their high implementation costs if compared to other available collocation methods, their presence in a general purpose software for IVPs is essentially missing [12]. A similar scenario may be depicted in the more recent branch of geometrical numerical integration, Gauss methods being among the first available symplectic Runge–Kutta methods.

In this paper we summarize the results presented in [1], where we introduce low–rank versions of Gauss collocation methods. More precisely, we recast their Butcher tableau by defining equivalent representations characterized by a rank–deficient array.

One appealing point that makes the study of such variants interesting is that, on the one hand, they define the very same numerical solution of the related classical Gauss method at each step of the integration procedure and, on the other hand, they open new routes of investigations for what concerns their implementation on a computer.

In particular, we here consider a special form of the methods where the so called *fundamental stages* are uncoupled. This leads to a nonlinear diagonal iteration for the solution of the associated nonlinear systems which may be carried out in parallel by constructing a number of processes equal to the number of fundamental stages (which, in turn, equals the number of stages of the underlying Gauss method).

For Gauss methods of order 4, 6 and 8, we show that formulae in this family exist such that the corresponding diagonal nonlinear iteration be *A*-convergent.

LOW–RANK GAUSS METHODS

Here we sketch the general idea to turn the Gauss collocation method with s stages (and hence of order $2s$) into an equivalent form with coefficients array A having dimension $k \geq s$. For a rigorous analysis about this approach we refer the reader to [1].

Our starting point is the nonlinear system defining the stages of the Gauss method applied to the initial value problem $y' = f(y)$, $y(t_0) = y_0 \in \mathbb{R}^m$, written with the aid of the W transformations [10]:

$$\widehat{Y} = e_s \otimes y_0 + h\left(\widehat{W} X_G \widehat{W}^T \widehat{\Omega}\right) \otimes I \cdot F(\widehat{Y}), \qquad (1)$$

where $\widehat{Y} = [\hat{y}_1^T, \dots, \hat{y}_s^T]^T$ is the vector containing the s stages of the method, $F(\widehat{Y}) = [f(\hat{y}_1)^T, \dots, f(\hat{y}_s)^T]^T$ contains the evaluations of the function f in the stages, I is the identity matrix of dimension m, e_s is the column vector of length s with unit entries, h is the step size of integration, $\widehat{\Omega}$ is the diagonal matrix whose diagonal entries are the weights $\hat{b}_1, \dots, \hat{b}_s$ of the Gauss-Legendre quadrature formula of order $2s$ defined at the Gauss nodes $\hat{c}_1 < \hat{c}_2 < \cdots < \hat{c}_s$ placed

Numerical Analysis and Applied Mathematics ICNAAM 2012
AIP Conf. Proc. 1479, 238-241 (2012); doi: 10.1063/1.4756107

on the interval $[0,1]$. Finally, the s-dimensional square matrices $\widehat{W}$ and X_G are

$$\widehat{W} = \begin{pmatrix} P_0(\hat{c}_1) & P_1(\hat{c}_1) & \dots & P_{s-1}(\hat{c}_1) \\ P_0(\hat{c}_2) & P_1(\hat{c}_2) & \dots & P_{s-1}(\hat{c}_2) \\ \vdots & \vdots & & \vdots \\ P_0(\hat{c}_s) & P_1(\hat{c}_s) & \dots & P_{s-1}(\hat{c}_s) \end{pmatrix}, \quad X_s = \begin{pmatrix} \frac{1}{2} & -\xi_1 & & \\ \xi_1 & 0 & \ddots & \\ & \ddots & \ddots & -\xi_{s-1} \\ & & \xi_{s-1} & 0 \end{pmatrix}, \quad \xi_j = \frac{1}{2\sqrt{(2j+1)(2j-1)}}, \tag{2}$$

where, for $j = 1, \dots, s$, $P_{j-1}(c)$ is the Legendre polynomial of degree $j-1$ shifted and normalized on the interval $[0,1]$, that is

$$\int_0^1 P_{i-1}(c) P_{j-1}(c) \mathrm{d}c = \begin{cases} 0, & \text{if } i \neq j \\ 1, & \text{if } i = j, \end{cases} \qquad i, j = 1, 2, \dots. \tag{3}$$

Let $\sigma_y : [0,1] \to \mathbb{R}^m$ be the polynomial of degree $s-1$ that interpolates the data $(t_0 + \hat{c}_i h, \hat{y}_i)$, namely $\sigma_y(t_0 + \hat{c}_i h) = \hat{y}_i$, $i = 1, \dots s$. We expand σ_y along the Legendre basis, that is $\sigma_y(t_0 + ch) = \sum_{j=1}^{s} \gamma_j P_{j-1}(c)$. The interpolation conditions become $\sum_{j=1}^{s} \gamma_j P_{j-1}(\hat{c}_i) = \hat{y}_i$, $i = 1, \dots, s$, which may be cast in matrix notation as $\widehat{Y} = (\widehat{W} \otimes I)\Gamma$, where Γ is the block-vector containing the vector coefficients γ_j.

Now let us consider k arbitrary abscissae $c_1, c_2, \dots, c_k$ on the interval $[0,1]$ and let $b_1, b_2, \dots, b_k$ the weights of the corresponding interpolation quadrature formula, whose degree of precision is assumed to be at least $2s-1$ (the degree of precision of the Gauss quadrature formula introduced above). For $i = 1, \dots k$, we set $y_i = \sigma_y(t_0 + c_i h) = \sum_{j=1}^{s} \gamma_j P_{j-1}(c_i)$ which, after introducing the block-vector $Y = [y_1^T, \dots, y_k^T]^T$ become $Y = (W \otimes I)\Gamma$ where W is the $k \times s$ matrix whose entries are $w_{ij} = P_{j-1}(c_i)$. Let $\Omega = \mathrm{diag}(b_1, \dots, b_k)$. Multiplying both sides of the last formula by $W^T \Omega \otimes I$ and considering that, by virtue of (3), $W^T \Omega W = I_k$ (the identity matrix of dimension k) yields $\Gamma = (W^T \Omega \otimes I) Y$. Therefore the block vectors $\widehat{Y}$ and Y are related as

$$\widehat{Y} = (\widehat{W} W^T \Omega \otimes I) Y. \tag{4}$$

We repeat the same computations as above by considering the polynomial $\sigma_f : [0,1] \to \mathbb{R}^m$ of degree $s-1$ that interpolates the data $(t_0 + \hat{c}_i h, f(\hat{y}_i))$, $i = 1, \dots, s$. Thus, after setting $Z = [\sigma_f(t_0 + c_1 h)^T, \dots, \sigma_f(t_0 + c_k h)^T]^T = (W \widehat{W}^T \widehat{\Omega} \otimes I) F(\widehat{Y})$ we obtain the analog of (4):

$$F(\widehat{Y}) = (\widehat{W} W^T \Omega \otimes I) Z. \tag{5}$$

Inserting (4) and (5) into (1) yields

$$(\widehat{W} W^T \Omega \otimes I) Y = e_s \otimes y_0 + h \left(\widehat{W} X_G W^T \Omega \otimes I \right) Z,$$

and multiplying both sides by $W \widehat{W}^{-1} \otimes I = W \widehat{W}^T \widehat{\Omega} \otimes I$ gives

$$Y = e_k \otimes y_0 + h \left(W X_G W^T \Omega \otimes I \right) Z. \tag{6}$$

In deriving (6) we have exploited the property that if V is any vector of length k containing the evaluations of a polynomial of degree at most $s-1$ (such as σ_y or σ_f) then $W W^T \Omega V = V$.

Summarizing, the original formulation of the nonlinear system (1) has been replaced by (6), where Z depends on Y via the formulae

$$Z = (W \widehat{W}^T \widehat{\Omega} \otimes I) F(\widehat{Y}) \qquad \text{and} \qquad \widehat{Y} = (\widehat{W} W^T \Omega \otimes I) Y. \tag{7}$$

Notice that $W X_G W^T$ in (6) has dimension k but rank s. This feature will be exploited in the next section to define a parallel iteration scheme to solve the nonlinear system in the unknown Y.

DIAGONALLY IMPLICIT GAUSS METHODS

Formulae (6)–(7) are a special case of a family of recently derived methods named Hamiltonian Boundary Value Methods (HBVMs) (see [3, 4, 5]). These may be viewed as Runge–Kutta methods with rank deficient coefficient matrix that are able to yield energy conservation if applied to a Hamiltonian system with polynomial Hamiltonian function. However, as is shown below, the solution of (6) is more conveniently computed by exploiting a HBVM rather than a Runge–Kutta formulation. Without going into details (see [1] for a complete analysis), we observe that matrix WX_G in (6) may be replaced by the $k \times s$ matrix $\mathscr{I}$ whose entries are $\mathscr{I}_{ij} = \int_0^{c_i} P_{j-1}(x)\mathrm{d}x$. After this substitution, the entries of the coefficient matrix $A \equiv \mathscr{I} W^T \Omega$ are

$$a_{ij} = b_j \sum_{\ell=1}^{s} P_\ell(c_j) \int_0^{c_i} P_\ell(x)\mathrm{d}x. \tag{8}$$

We can further generalize our method by allowing different quadrature formulae (and hence weights b_j) on each row of matrix A. Our aim is to make some coefficients a_{ij} vanish. More precisely, let us pick s of the k abscissae c_i and assume, without loss of generality, that they are $c_1, c_2, \dots, c_s$. In defining the first stage y_1, we will consider the quadrature weights $b_j^{(1)}$ corresponding to the abscissae $c_1, c_{s+1}, c_{s+2}, \dots, c_k$. From (8) it will follow that $a_{12} = \cdots = a_{1s} = 0$. Analogously, for $i = 2, \dots, s$, we define the coefficients $b_j^{(i)}$, involved in the definition of the i-th row of matrix A, as the weights of the quadrature formula corresponding to the abscissae $c_i, c_{s+1}, c_{s+2}, \dots, c_k$. In so doing, the leading principal submatrix of dimension s of A has a diagonal structure and will be denoted by D.

That the rank of matrix A is s rather than k has an important consequence: $k - s$ stages among the y_i are linear combinations of y_0 and the remaining stages. According to the theory on HBVMs, to conveniently exploit this property we split the vector Y in two parts, namely $Y = [Y_f^T, Y_s^T]^T$, where $Y_f = [y_1^T, \dots, y_s^T]^T$ contains the first s stages and $Y_s = [y_{s+1}^T, \dots, y_k^T]^T$ contains the remaining $k-s$ stages. Then we can express the latter vector as $Y_s = a \otimes y_0 + (C \otimes I)Y_f$, where a is a column vector of length $k-s$ and C is a matrix of size $(k-s) \times s$. For this reason the stages belonging to Y_f are referred to as *fundamental stages* while those belonging to Y_s are called *silent stages* (they are not to be regarded as unknowns).

The above argument shows that the nonlinear system (6) is equivalent to

$$\begin{aligned} Y_f &= e_s \otimes y_0 + h(D \otimes I)Z_f + h(B \otimes I)Z_s, \\ Y_s &= a \otimes y_0 + (C \otimes I)Y_f, \end{aligned} \tag{9}$$

where Z_f and Z_s are obtained from Z in the obvious way.

Remark 1 *In the special case where $c_i = \hat{c}_i$, $i = 1, \dots, s$, from* (7) *we get $Y_f = \widehat{Y}$ and $Z_f = F(\widehat{Y})$, and the resulting method has a simplified formulation.*

Remark 2 *Replacing in* (9) *Z_f by $F(Y_f)$ and Z_s by $F(Y_s)$, result in the general class of Diagonally Implicit HBVMs. When applied to autonomous Hamiltonian systems, these methods fail to be symplectic but can conserve the energy function if k is large enough [2].*

DIAGONAL ITERATION AND LINEAR CONVERGENCE ANALYSIS

System (9) naturally suggests the following diagonal nonlinear iteration to determine its solution:

$$Y_f^{(k+1)} - h(D \otimes I)Z_f^{(k+1)} = e_s \otimes y_0 + h(B \otimes I)Z_s^{(k)}. \tag{10}$$

Here we are viewing Z_f as a function of Y_f. Actually, since Z_f is nothing but the evaluation, at the abscissae $c_1, \dots, c_s$, of the polynomial that interpolates $F(\widehat{Y})$ over the Gauss-Legendre nodes, we deduce that

$$Z_f = F(Y_f) + O(h). \tag{11}$$

We apply a single step of a simplified Newton method to get an approximation of $Y_f^{(k+1)}$ and then proceed with the primary iteration (10) until we get convergence. In view of (11), which allows us to replace Z_f by $F(Y_f)$ thus neglecting the $O(h^2)$ term in the approximation of the Jacobian, the resulting iteration scheme reads:

$$\begin{aligned} &(I_{ms} - h(D \otimes I)J)\,\Delta^{(k)} = Y_f^{(k)} - e_s \otimes y_0 - h(D \otimes I)Z_f^{(k)} - h(B \otimes I)Z_s^{(k)}, \\ &Y_f^{(k+1)} = Y_f^{(k)} + \Delta^{(k)}. \end{aligned} \tag{12}$$

where $J = \mathrm{diag}(J_1, \dots, J_s)$ ia a block diagonal matrix containing approximations of the Jacobian of f at the points $y_1, \dots, y_s$.The simplest choice is $J_i = J_0 = \frac{\partial f}{\partial y}(y_0)$ which turns out to be appropriate when the process is handled by a single thread.

Due to the A-stability property of Gauss methods, it is desirable that the scheme be convergent whenever the method is applied to the scalar test equation $y' = \lambda y$ and $\lambda < 0$. If so, the iteration is said A-convergent [13, 14**?**]. Setting $q = h\lambda$ and considering that in the linear case $Z = Y$, the iteration matrix $M(q)$ associated to the scheme (10) is

$$M(q) = q(I - qD)^{-1}BC.$$

In the case where D has positive entries, A-convergence is equivalent to the requirement $\rho^* \equiv \max_{x\in\mathbb{R}^+} \rho(M(ix)) \leq 1$, where $\rho(M)$ denotes the spectral radius of the matrix M and i is the imaginary unit. Another desirable property is that $\rho(q) \to 0, \qquad \text{as} \quad q \to \infty$. In this case fast convergence is guaranteed in correspondence to large values of $|q|$ in the left half complex plane and consequently the iteration is very suited for stiff problems [13]. At infinity the iteration matrix $M(q)$ assumes the form $M(\infty) = -D^{-1}BC$ and consequently $\rho_{\mathrm{inf}} \equiv \lim_{q\to\infty} M(q) = \rho(D^{-1}BC)$. On the other hand, when $q \simeq 0$ (non-stiff problems) we have $M(q) \simeq qBC$ and the asymptotic behavior is governed by $\rho_0 \equiv \rho(BC)$. The parameters ρ_{inf} and ρ_0 are referred to as *stiff* and *non-stiff amplification factors*.

The freedom in the choice of the k abscissae $c_1, \dots, c_k$ has then been exploited to improve as much as possible the above parameters. For the methods of order 4, 6 and 8 we have chosen $k = 6, 9, 12$ respectively, and using minimization technique we have found abscissae distributions that made the iteration associated with the three methods A-convergent. Table 1 summarizes the linear convergence properties of the three schemes.

TABLE 1. Convergence parameters for Gauss methods of order 4, 6 and 8.

method	ρ^*	ρ_{inf}	ρ_0
order 4	$1.56 \cdot 10^{-1}$	$1.43 \cdot 10^{-9}$	$9.30 \cdot 10^{-2}$
order 6	$5.36 \cdot 10^{-1}$	$1.16 \cdot 10^{-3}$	$3.23 \cdot 10^{-1}$
order 8	$4.35 \cdot 10^{-1}$	$1.77 \cdot 10^{-2}$	$8.27 \cdot 10^{-1}$

REFERENCES

1. L. Brugnano and F. Iavernaro. *A parallel implementation of Gauss methods for ODEs*, in preparation.
2. L. Brugnano and F. Iavernaro. *Diagonally implicit low-rank Runge–Kutta methods of collocation type*, in preparation.
3. L. Brugnano, F. Iavernaro and D. Trigiante. *The Hamiltonian BVMs (HBVMs) Homepage*, `http://web.math.unifi.it/users/brugnano/HBVM/`.
4. L. Brugnano, F. Iavernaro and D. Trigiante. Hamiltonian Boundary Value Methods (Energy Preserving Discrete Line Methods). *Journal of Numerical Analysis, Industrial and Applied Mathematics* **5**, 1-2 (2010) 17–37.
5. L. Brugnano, F. Iavernaro and D. Trigiante. The Lack of Continuity and the Role of Infinite and Infinitesimal in Numerical Methods for ODEs: the Case of Symplecticity. *Applied Mathematics and Computation* **218** (2012) 8053–8063.
6. L. Brugnano, C. Magherini. Recent Advances in Linear Analysis of Convergence for Splittings for Solving ODE problems. *Applied Numerical Mathematics* **59** (2009) 542–557.
7. J.C. Butcher. *Numerical Methods for Ordinary Differetial Equations*, Second Edition, Wiley, 2008.
8. Feng Kang and Qin Meng-zhao. The symplectic methods for the computation of Hamiltonian equations. *Lecture Notes in Math.* **1297** (1987) 1–37.
9. E. Hairer, C. Lubich and G. Wanner. *Geometric Numerical Integration. Structure-Preserving Algorithms for Ordinary Differential Equations*, Second edition, Springer, Berlin, 2006.
10. E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II. Stiff and Differential- Algebraic Problems*, Second ed., Springer Series in Computational Mathematics 14, Springer-Verlag Berlin, 1996.
11. B. Leimkuhler and S. Reich. *Simulating Hamiltonian Dynamics*. Cambridge University Press, Cambridge, 2004.
12. F. Mazzia and C. Magherini, *Test Set for Initial Value Problem Solvers*, release 2.4, February 2008, Department of Mathematics, University of Bari and INdAM, Research Unit of Bari, available at `http://www.dm.uniba.it/~testset`.
13. P. J. van der Houwen, J. J. B. de Swart. *Triangularly Implicit Iteration Methods for ODE-IVP Solvers*, SIAM J. Sci. Comput., 18, 1997, pp. 41-55.
14. P. J. van der Houwen, J. J. B. de Swart. *Parallel Linear System Solvers for Runge-Kutta Methods*, Adv. Comput. Math., 7,1-2, 1997, pp. 157-181.